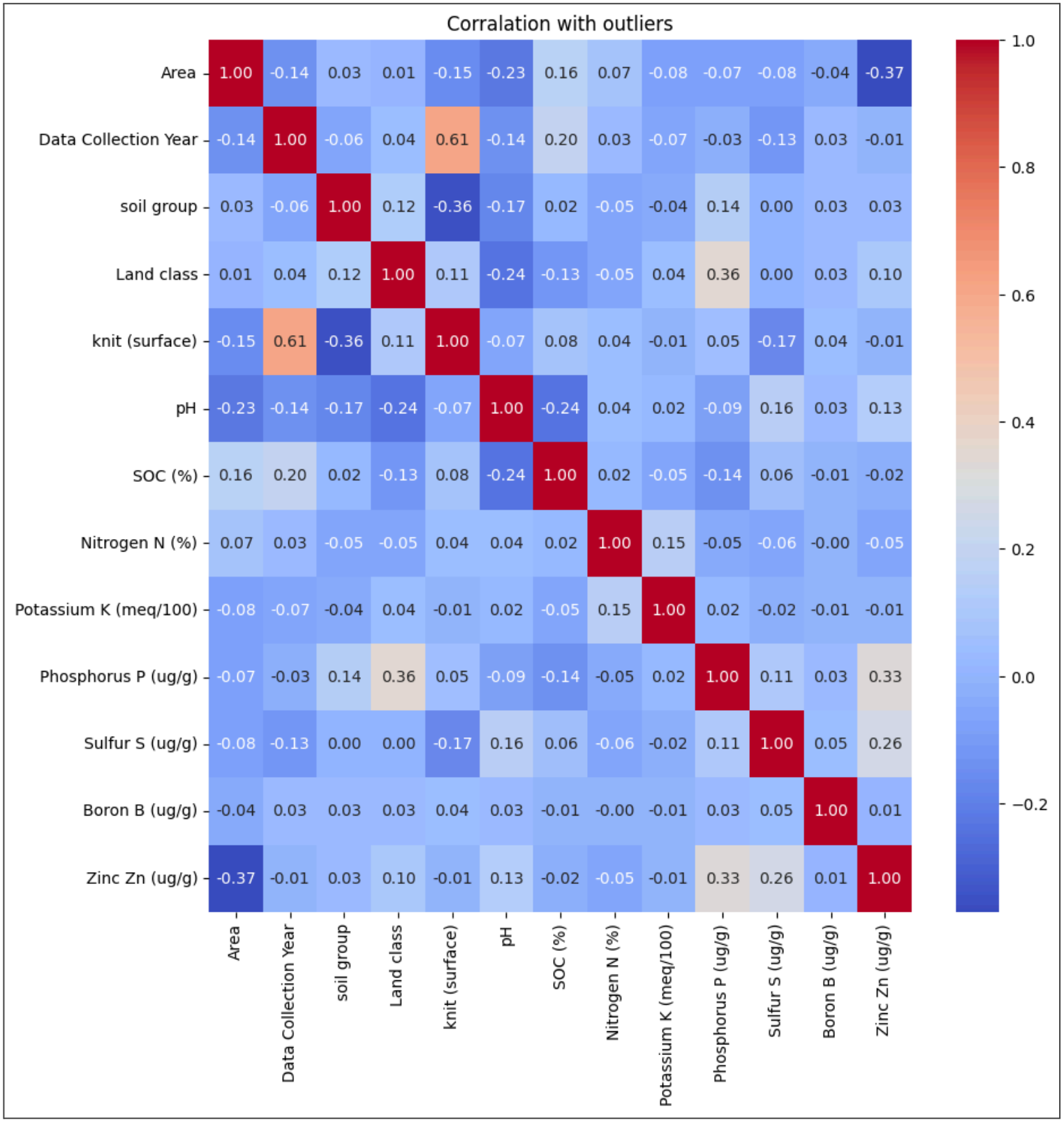
EDA (Basic)

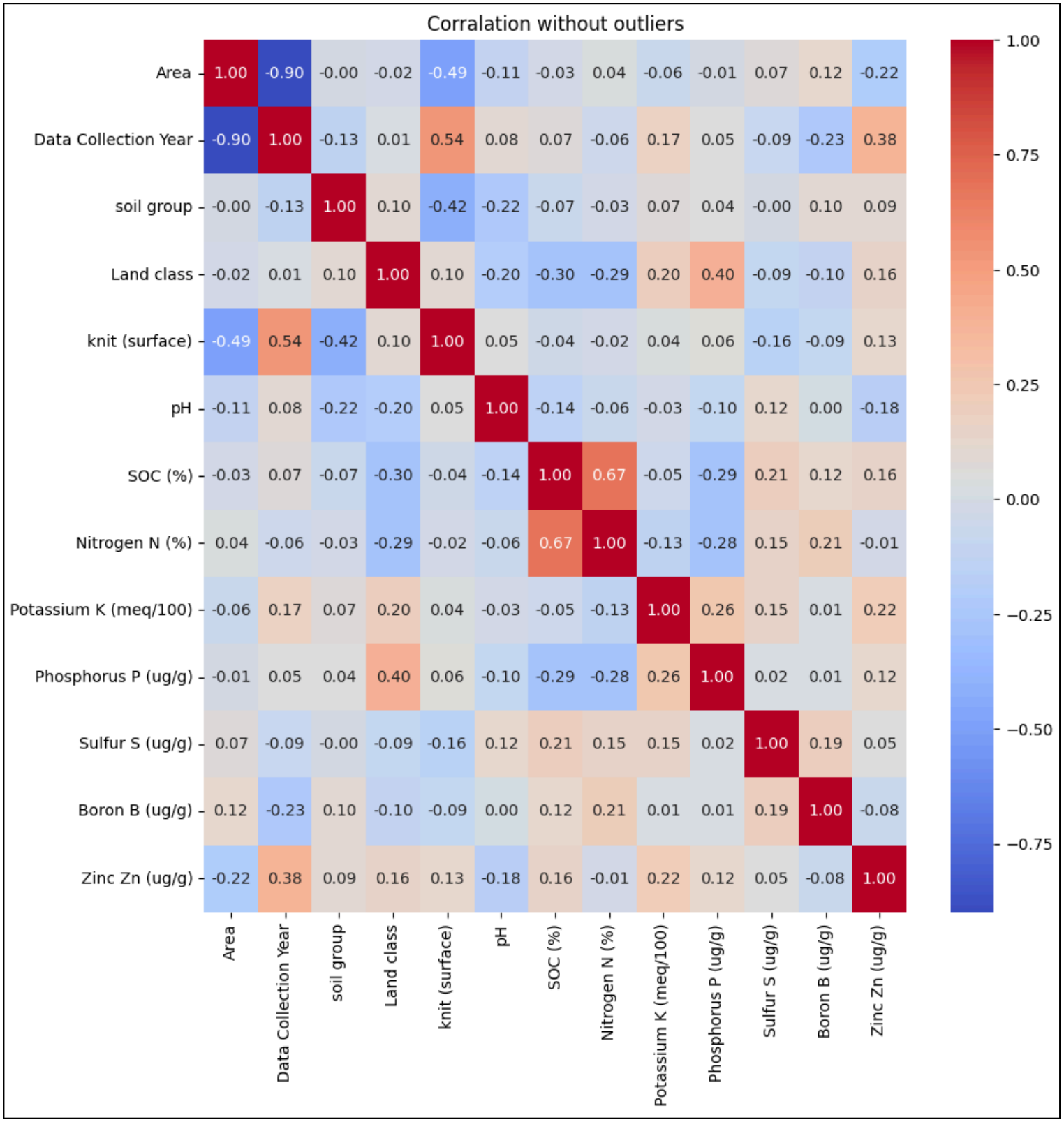
Element	Replacement	Replacement Action
Zinc Zn (ug/g)	Tuckey	Replace with group median
Boron B (ug/g)	Tuckey	Replace with group median
Sulfur S (ug/g)	Tuckey	Replace with group median
Phosphorus P (ug/g)	Tuckey	Replace with group median
Potassium K (meq/100)	Tuckey	Replace with group median
Nitrogen N (%)	Tuckey	Replace with group median
SOC (%)	Tuckey	Replace with group median
pH	Tuckey	Replace with group median

Notebook Link →

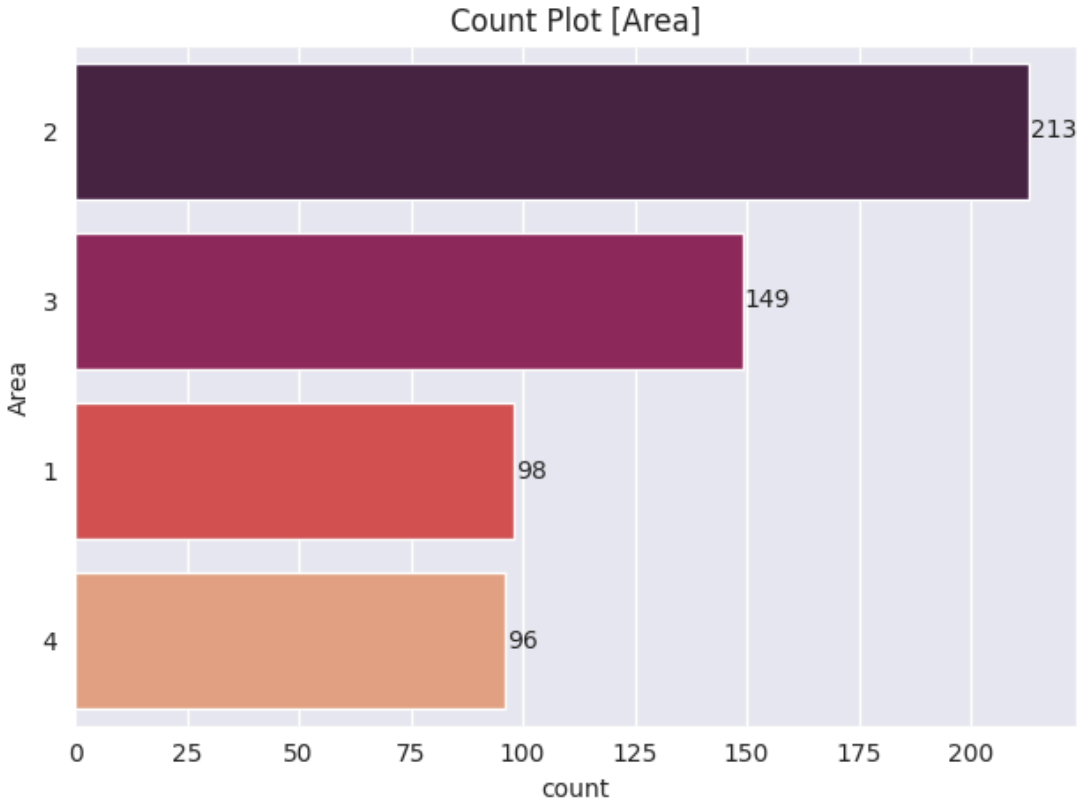
Knit (surface)

loam	358	NA	157	Clay loam	76
loam clay	15	brick	9	in the sand	3


Corralation with outliers
Area 1.00 -0.14 0.03 0.01 -0.15 -0.23 0.16 0.07 -0.08 -0.07 -0.08 -0.04 -0.37
Data Collection Year -0.14 1.00 -0.06 0.04 0.61 -0.14 0.20 0.03 -0.07 -0.03 -0.13 0.03 -0.01
soil group 0.03 -0.06 1.00 0.12 -0.36 -0.17 0.02 -0.05 -0.04 0.14 0.00 0.03 0.03
Land class 0.01 0.04 0.12 1.00 0.11 -0.24 -0.13 -0.05 0.04 0.36 0.00 0.03 0.10
knit (surface) -0.15 0.61 -0.36 0.11 1.00 -0.07 0.08 0.04 -0.01 0.05 -0.17 0.04 -0.01
pH -0.23 -0.14 -0.17 -0.24 -0.07 1.00 -0.24 0.04 0.02 -0.09 0.16 0.03 0.13
SOC (%) 0.16 0.20 0.02 -0.13 0.08 -0.24 1.00 0.02 -0.05 -0.14 0.06 -0.01 -0.02
Nitrogen N (%) 0.07 0.03 -0.05 -0.05 0.04 0.04 0.02 1.00 0.15 -0.05 -0.06 -0.00 -0.05
Potassium K (meq/100) -0.08 -0.07 -0.04 0.04 -0.01 0.02 -0.05 0.15 1.00 0.02 -0.02 -0.01 -0.01
Phosphorus P (ug/g) -0.07 -0.03 0.14 0.36 0.05 -0.09 -0.14 -0.05 0.02 1.00 0.11 0.03 0.33
Sulfur S (ug/g) -0.08 -0.13 0.00 0.00 -0.17 0.16 0.06 -0.06 -0.02 0.11 1.00 0.05 0.26
Boron B (ug/g) -0.04 0.03 0.03 0.03 0.04 0.03 -0.01 -0.00 -0.01 0.03 0.05 1.00 0.01
Zinc Zn (ug/g) -0.37 -0.01 0.03 0.10 -0.01 0.13 -0.02 -0.05 -0.01 0.33 0.26 0.01 1.00
Area
Data Collection Year
soil group
Land class
knit (surface)
pH
SOC (%)
Nitrogen N (%)
Potassium K (meq/100)
Phosphorus P (ug/g)
Sulfur S (ug/g)
Boron B (ug/g)
Zinc Zn (ug/g)
1.0
0.8
0.6
0.4
0.2
0.0
−0.2

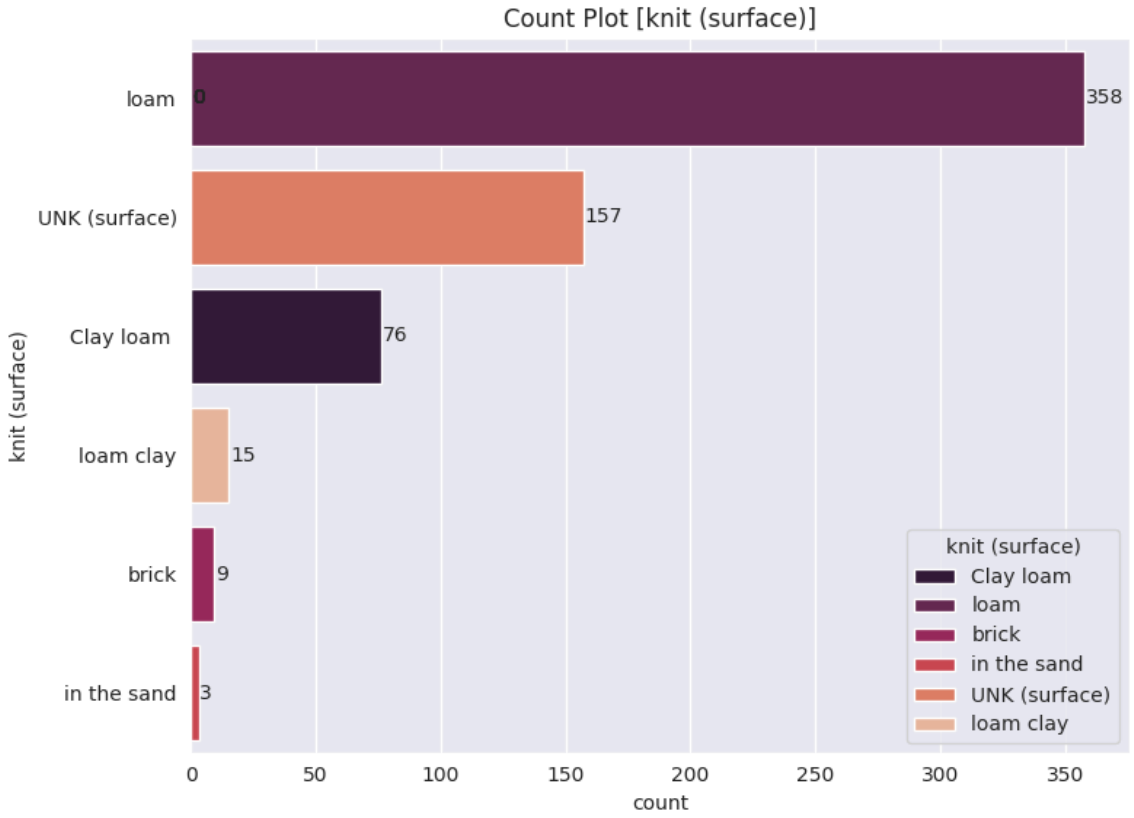

Corralation without outliers
Area 1.00 -0.90 -0.00 -0.02 -0.49 -0.11 -0.03 0.04 -0.06 -0.01 0.07 0.12 -0.22
Data Collection Year -0.90 1.00 -0.13 0.01 0.54 0.08 0.07 -0.06 0.17 0.05 -0.09 -0.23 0.38
soil group -0.00 -0.13 1.00 0.10 -0.42 -0.22 -0.07 -0.03 0.07 0.04 -0.00 0.10 0.09
Land class -0.02 0.01 0.10 1.00 0.10 -0.20 -0.30 -0.29 0.20 0.40 -0.09 -0.10 0.16
knit (surface) -0.49 0.54 -0.42 0.10 1.00 0.05 -0.04 -0.02 0.04 0.06 -0.16 -0.09 0.13
pH -0.11 0.08 -0.22 -0.20 0.05 1.00 -0.14 -0.06 -0.03 -0.10 0.12 0.00 -0.18
SOC (%) -0.03 0.07 -0.07 -0.30 -0.04 -0.14 1.00 0.67 -0.05 -0.29 0.21 0.12 0.16
Nitrogen N (%) 0.04 -0.06 -0.03 -0.29 -0.02 -0.06 0.67 1.00 -0.13 -0.28 0.15 0.21 -0.01
Potassium K (meq/100) -0.06 0.17 0.07 0.20 0.04 -0.03 -0.05 -0.13 1.00 0.26 0.15 0.01 0.22
Phosphorus P (ug/g) -0.01 0.05 0.04 0.40 0.06 -0.10 -0.29 -0.28 0.26 1.00 0.02 0.01 0.12
Sulfur S (ug/g) 0.07 -0.09 -0.00 -0.09 -0.16 0.12 0.21 0.15 0.15 0.02 1.00 0.19 0.05
Boron B (ug/g) 0.12 -0.23 0.10 -0.10 -0.09 0.00 0.12 0.21 0.01 0.01 0.19 1.00 -0.08
Zinc Zn (ug/g) -0.22 0.38 0.09 0.16 0.13 -0.18 0.16 -0.01 0.22 0.12 0.05 -0.08 1.00
Area
Data Collection Year
soil group
Land class
knit (surface)
pH
SOC (%)
Nitrogen N (%)
Potassium K (meq/100)
Phosphorus P (ug/g)
Sulfur S (ug/g)
Boron B (ug/g)
Zinc Zn (ug/g)
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75

Count Plot (Area)



- Most values are from **Mithapukur** (213) and the least are from **Gangachara** (96)

Count Plot (Knit (surface))



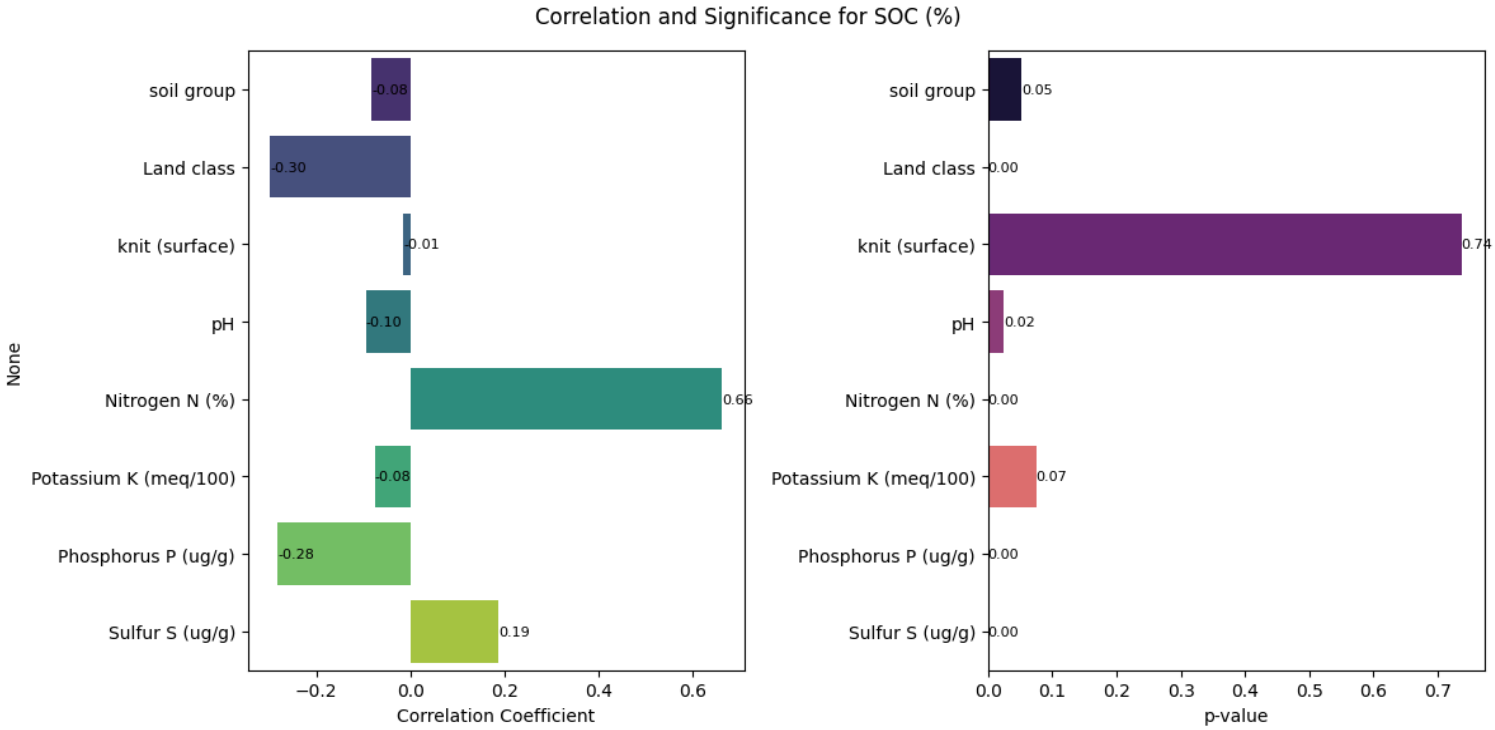
- **157** surface type without names replace with UNK (surface)
- Are **Clay Loam** & **Loam Clay** similar? If they are we can categorize them as one
    - Clay Loam: Contains about 20-30% clay, with a more balanced loamy texture.
    - Loam Clay: Contains over 30% clay, giving it a denser, more clayey texture.
- **Brick** & **In the Sand** has too few values

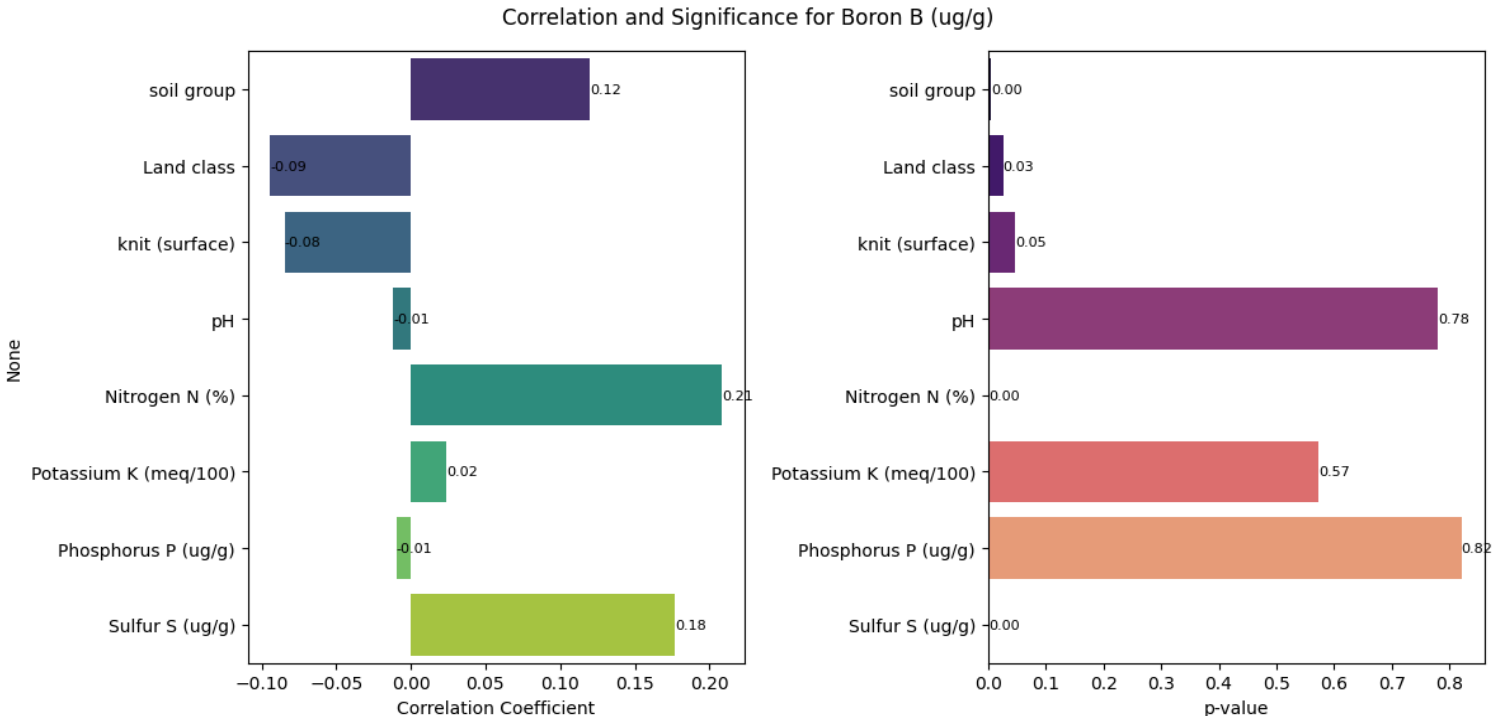
Skewness

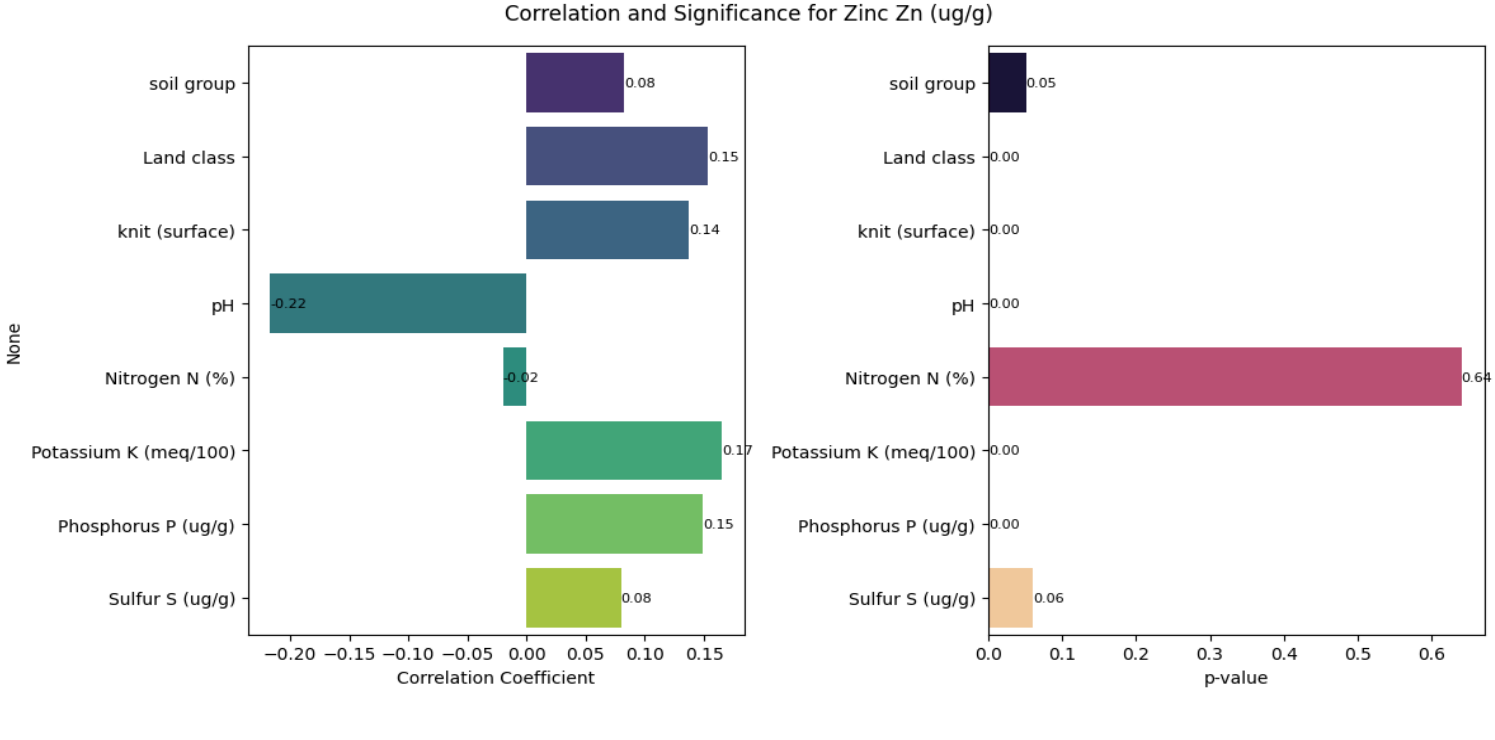
	Skewness	
Element	**With Outliers**	**Without Outliers**
Zinc Zn (ug/g)	3.95	0.37
Boron B (ug/g)	3.43	0.72
Sulfur S (ug/g)	1.90	0.45
Phosphorus P (ug/g)	2.62	1.14
Potassium K (meq/100)	22.49	1.02
Nitrogen N (%)	24.27	0.32
SOC (%)	9.63	0.37
pH	1.25	0.30

Correlation and p-value Between Target and Feature variables

Element		Soil Group	Land class	Knit (surface)	pH	Nitrogen N (%)	Potassium K (meq/100)	Phosphorus P (ug/g)	Sulfur S (ug/g)
Boron B (ug/g)	Corr	0.120	-0.094	-0.084	-0.012	0.209	0.024	-0.010	0.177
	p-value	0.0046	0.0263	0.046	0.779	0.0000	0.5725	0.8204	0.0000
SOC (%)	Corr	-0.082	-0.297	-0.014	-0.095	0.662	-0.076	-0.282	0.187
	p-value	0.0524	0.000	0.7366	0.025	0.000	0.075	0.000	0.000
Zinc Zn (ug/g)	Corr	0.083	0.153	0.137	-0.217	-0.020	0.165	0.149	0.080
	p-value	0.0513	0.0003	0.0012	0.000	0.6405	0.0001	0.0004	0.0600

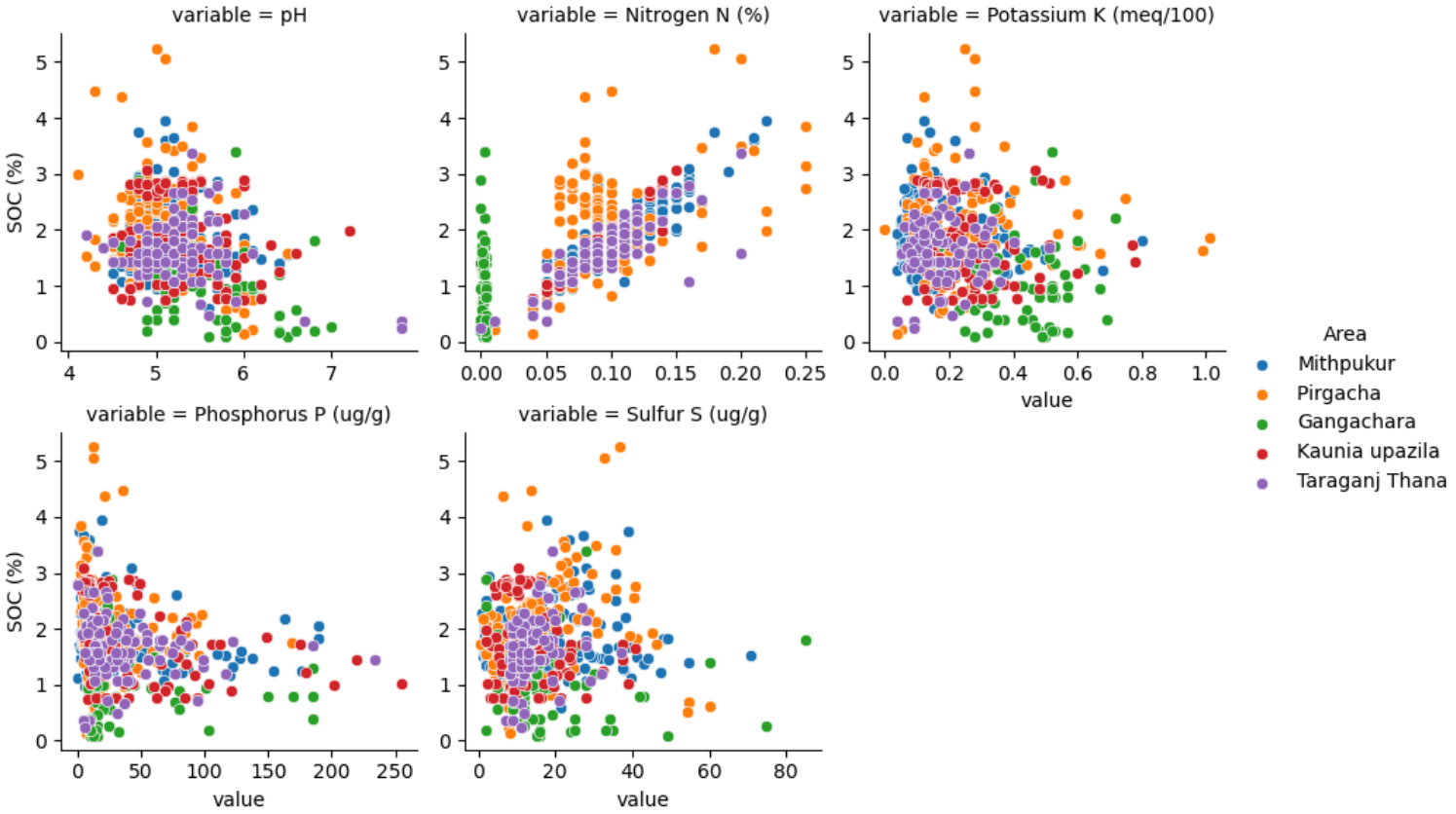

Correlation and Significance for SOC (%)
soil group
Land class
knit (surface)
pH
Nitrogen N (%)
Potassium K (meq/100)
Phosphorus P (ug/g)
Sulfur S (ug/g)
None
-0.08
-0.30
-0.01
-0.10
0.66
-0.08
-0.28
0.19
−0.2
0.0
0.2
0.4
0.6
Correlation Coefficient
0.05
0.00
0.74
0.02
0.00
0.07
0.00
0.00
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
p-value


Correlation and Significance for Boron B (ug/g)
soil group
Land class
knit (surface)
pH
Nitrogen N (%)
Potassium K (meq/100)
Phosphorus P (ug/g)
Sulfur S (ug/g)
None
0.12
-0.09
-0.08
-0.01
0.21
0.02
-0.01
0.18
−0.10
−0.05
0.00
0.05
0.10
0.15
0.20
Correlation Coefficient
0.00
0.03
0.05
0.78
0.00
0.57
0.82
0.00
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
p-value

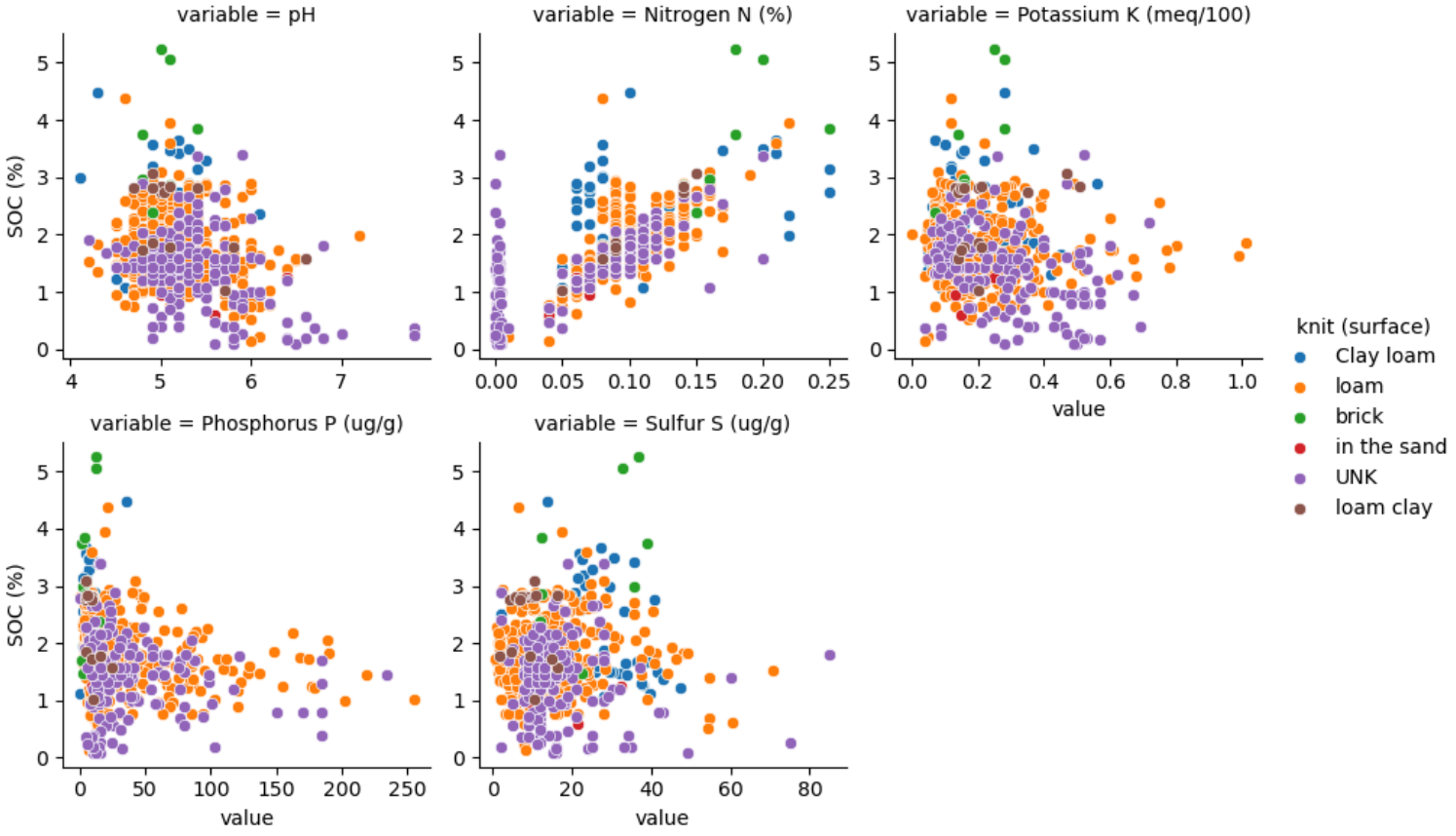

Correlation and Significance for Zinc Zn (ug/g)
soil group
Land class
knit (surface)
pH
Nitrogen N (%)
Potassium K (meq/100)
Phosphorus P (ug/g)
Sulfur S (ug/g)
None
0.08
0.15
0.14
-0.22
-0.02
0.17
0.15
0.08
−0.20
−0.15
−0.10
−0.05
0.00
0.05
0.10
0.15
Correlation Coefficient
0.05
0.00
0.00
0.00
0.64
0.00
0.00
0.06
0.0
0.1
0.2
0.3
0.4
0.5
0.6
p-value

SOC %

- Area Wise

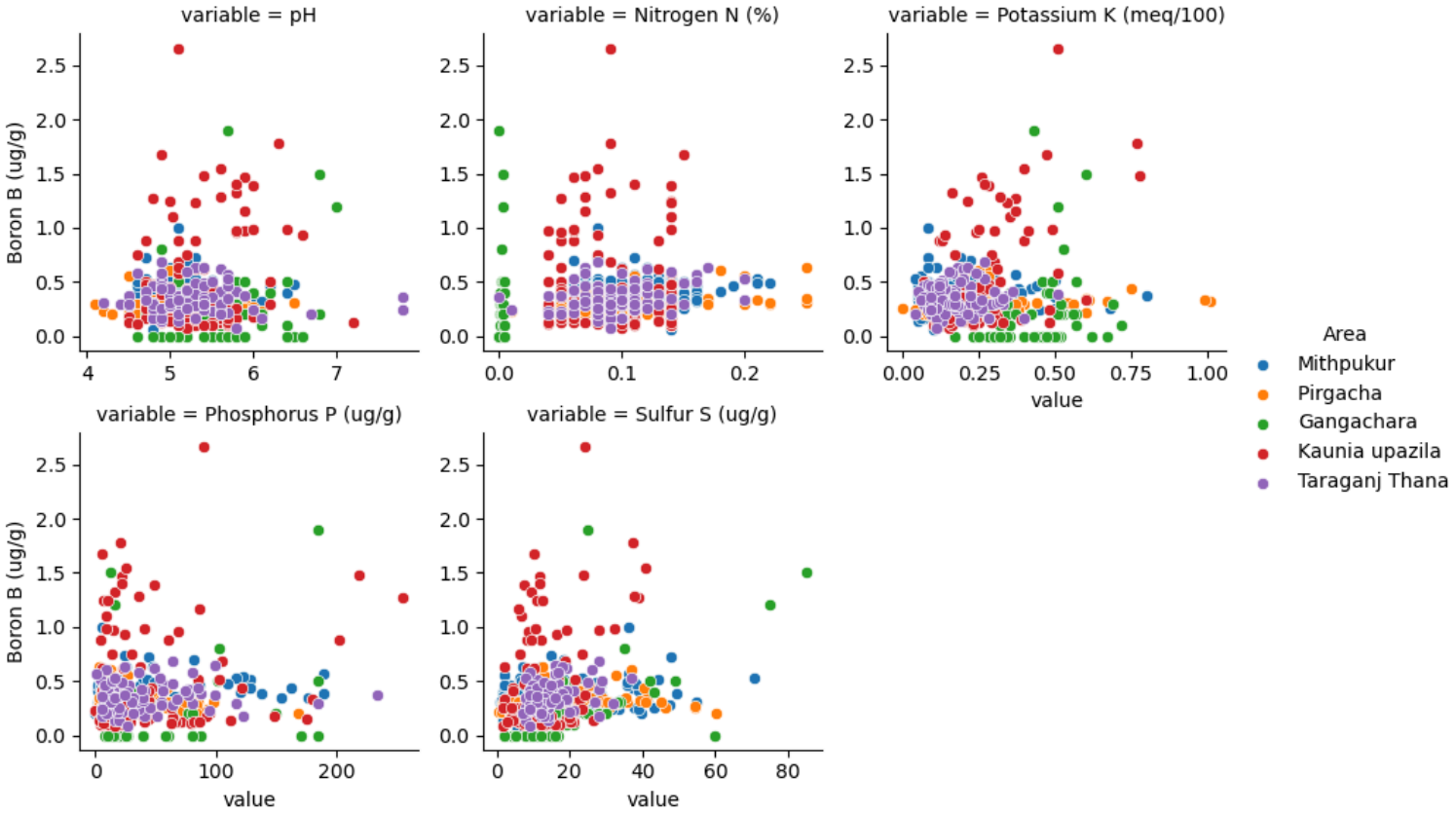


- Knit (surface)

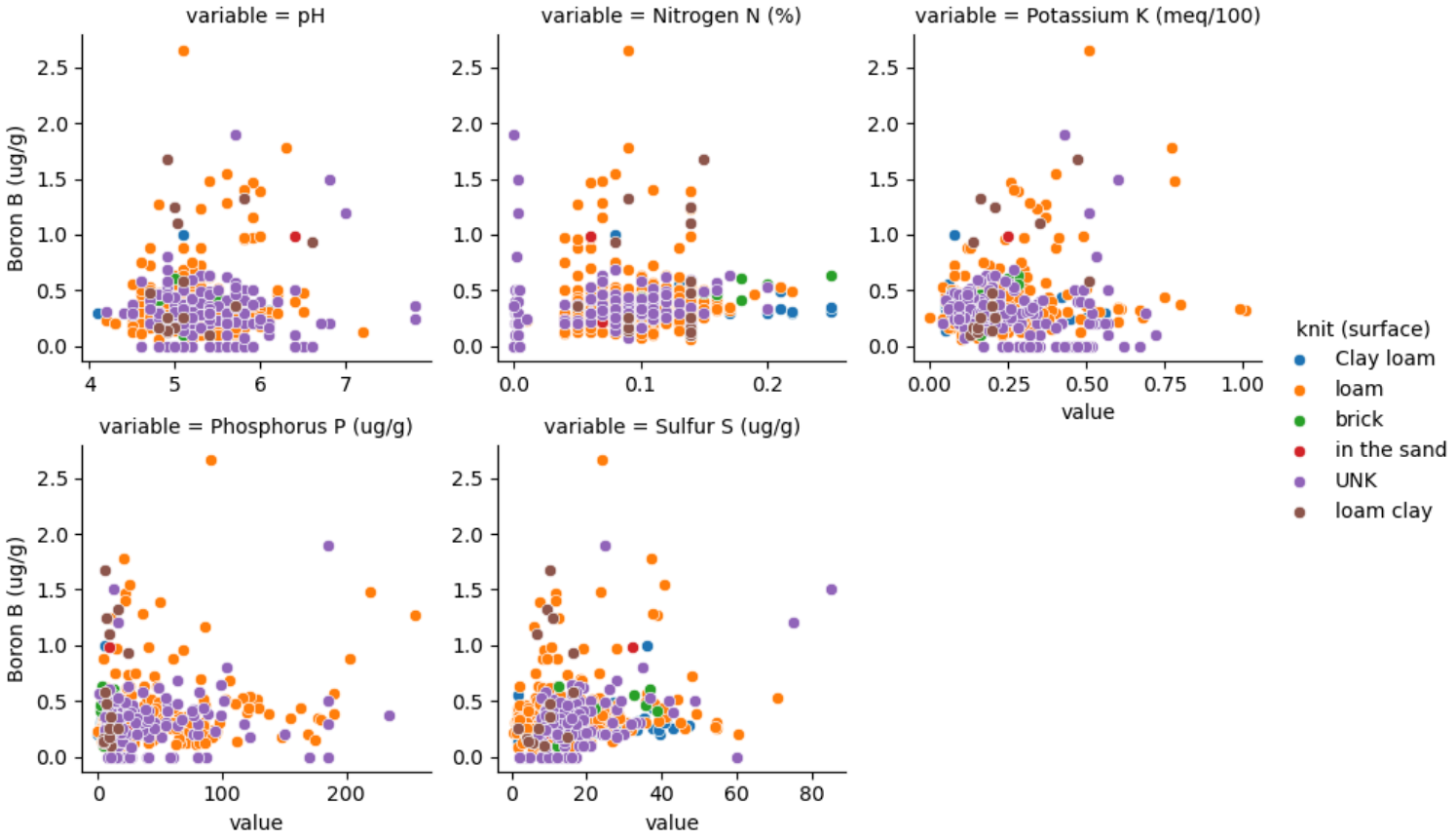


Boron B

- Area Wise

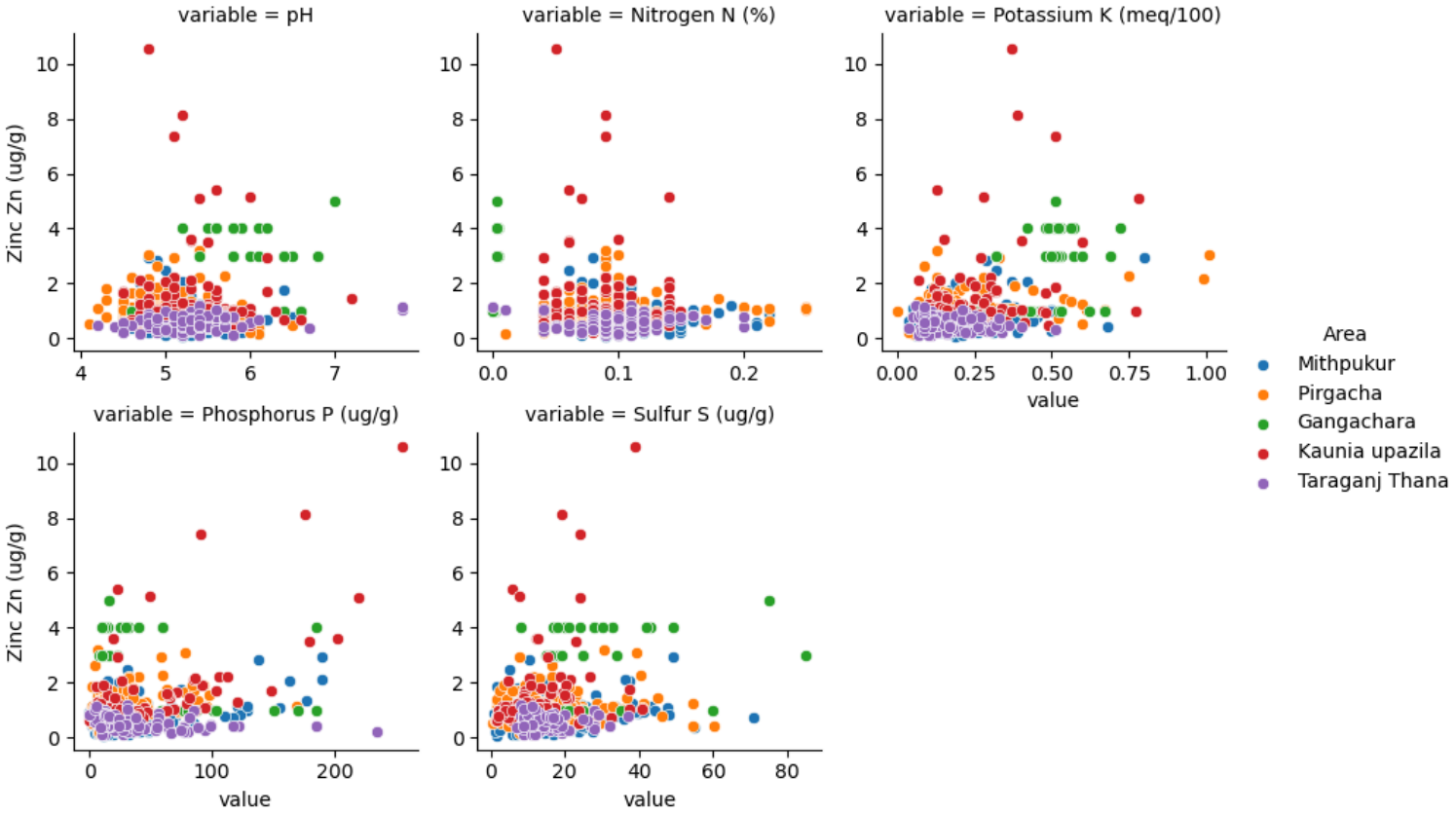


- Knit (surface)

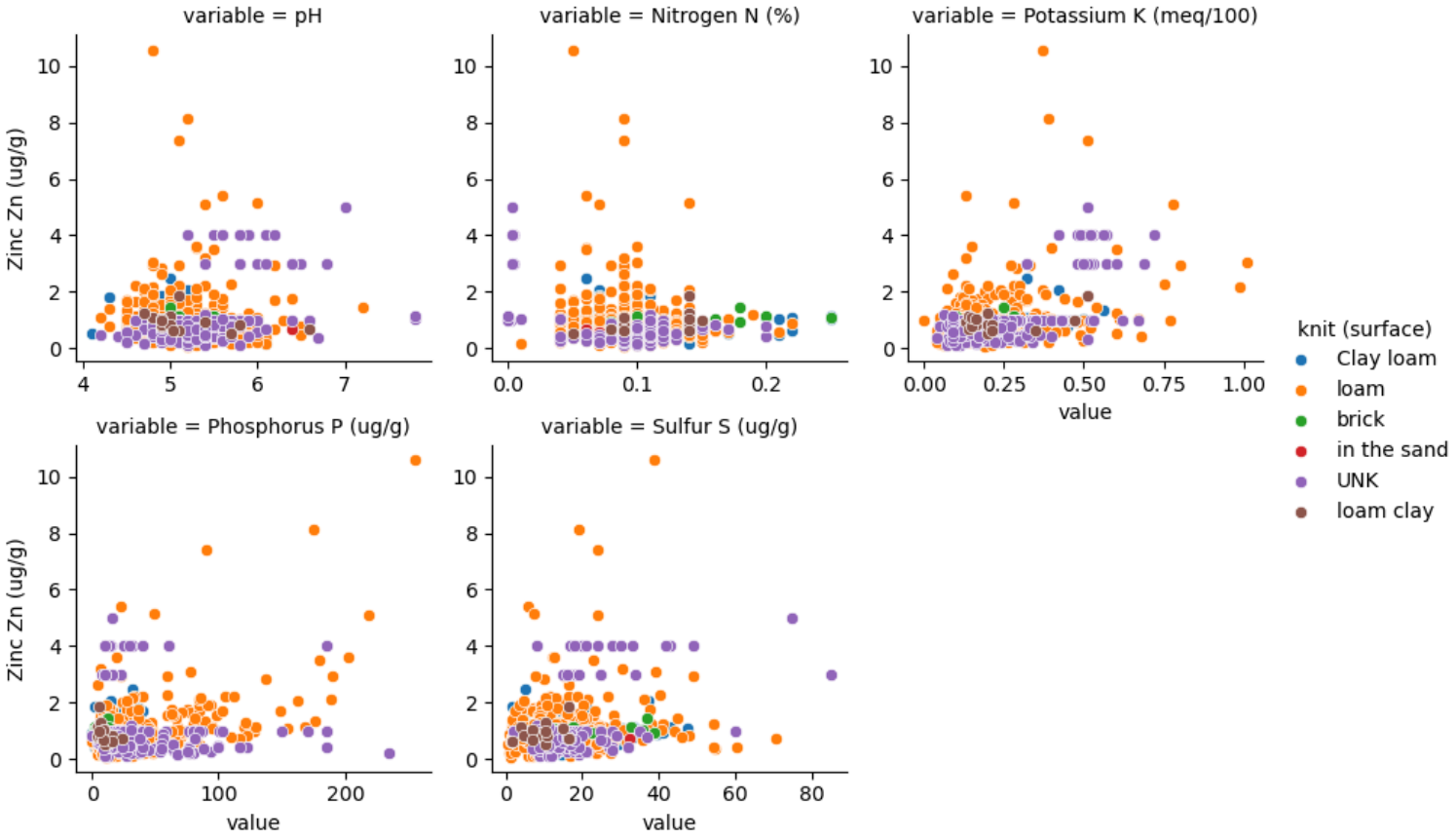


Zinc Zn

- Area Wise



- Knit (surface)



Preprocessed DataFrame

	count	mean	std	min	25%	50%	75%	max
Area	556.0	2.437050	0.972409	1.00	2.00	2.000	3.0000	4.00
Data Collection Year	556.0	2005.557554	5.667858	1997.00	2005.00	2005.000	2005.0000	2016.00
soil group	556.0	9.530576	6.800435	0.00	5.00	5.000	15.0000	26.00
Land class	556.0	2.372302	1.472837	0.00	1.00	1.000	4.0000	4.00
knit (surface)	556.0	2.924460	1.638956	0.00	1.00	4.000	4.0000	5.00
pH	556.0	5.159856	0.358771	4.20	4.90	5.100	5.4000	6.10
SOC (%)	556.0	1.809942	0.560554	0.23	1.44	1.755	2.1600	3.30
Nitrogen N (%)	556.0	0.095849	0.028638	0.04	0.08	0.090	0.1100	0.17
Potassium K (meq/100)	556.0	0.181577	0.097745	0.00	0.11	0.160	0.2400	0.52
Phosphorus P (ug/g)	556.0	20.194392	15.424684	0.06	8.10	15.495	26.4700	64.00
Sulfur S (ug/g)	556.0	12.513133	6.437698	0.40	8.30	12.000	16.3000	29.30
Boron B (ug/g)	556.0	0.308520	0.102483	0.06	0.24	0.300	0.3600	0.58
Zinc Zn (ug/g)	556.0	0.771955	0.351437	0.08	0.50	0.750	1.0225	1.72